

Recopilación y Recomendación de Objetos de Aprendizaje

Claudia Deco^{1,2}, Ana Casali^{1,3}, Cristina Bender^{1,2}

¹ Facultad de Ciencias Exactas, Ingeniería y Agrimensura,
Universidad Nacional de Rosario, Rosario, Argentina.
{deco, acasali, bender}@fceia.unr.edu.ar

² Departamento de Investigación Institucional, Facultad de Química e
Ingeniería, Universidad Católica Argentina Campus Rosario, Argentina

³ Centro Internacional Franco Argentino de Ciencias de la Información y de
Sistemas CIFASIS

Resumen

En este trabajo se presenta una línea de investigación orientada hacia el desarrollo de distintas herramientas informáticas que ayuden a la gestión de repositorios de objetos de aprendizaje. En este sentido se trabaja en dos aspectos fundamentales para mejorar su usabilidad. Un aspecto es dar soporte a las tareas de recopilación de documentos que realiza el administrador del repositorio con el objetivo de detectar documentos plausibles de ser cargados en estos repositorios juntos con sus metadatos de interés. El otro aspecto es continuar trabajando en la recomendación de objetos de aprendizaje considerando no sólo el perfil del usuario y el diseño instruccional del material educativo sino también la valoración colaborativa de grupos de estudiantes similares.

Palabras clave: Objetos de aprendizaje, Repositorios, Recopilación de Información, Extracción de Metadatos, Sistemas Recomendadores.

Contexto

Esta línea de I+D se está llevando a cabo a través de varios proyectos de investigación:

- Red Cyted: RIURE Red Iberoamericana para la Usabilidad de

Repositorios Educativos (2013-2016). Universidades participantes: Universidad Nacional de Rosario, Argentina; Fundação Universidade Federal do Pampa, Brasil; Escuela Superior Politécnica del Litoral, Ecuador; Universidad Complutense de Madrid, España; Instituto Superior de Engenharia do Porto, Portugal; y Universidad de la República, Uruguay. Representantes por UNR: Ana Casali, Claudia Deco.

- PID UNR: Sistemas Multiagentes y sus Aplicaciones a la Gestión de Objetos Digitales Educativos [ING466] dirigido por Ana Casali (2014-2017).

- PID UNR: Búsqueda Personalizada de Recursos Educativos. [ING491] dirigido por Claudia Deco (2015-2018). En evaluación. Continuidad del PID UNR [ING348] (2011-2014).

- Proyecto UCA: Mejora de la enseñanza de las ciencias en las carreras de Ingeniería mediante la búsqueda personalizada de recursos educativos. Claudia Deco. (2014-2015).

Introducción

En los últimos años, el desarrollo de repositorios institucionales de acceso abierto ha sido un tema prioritario en las políticas de educación, ciencia y técnica de muchos países y en particular, la

creación de repositorios de objetos digitales educativos en las universidades públicas de Argentina, es una prioridad en el marco de las políticas de nuestro país. Estos repositorios tienen como objetivo la preservación y diseminación de conocimiento abierto, gratuito y accesible para toda la ciudadanía. En este sentido, se consideran como Objetos Digitales Educativos (ODEs) a almacenarse en estos repositorios, a toda la producción resultante de la actividad docente, científica, tecnológica y artística de la institución, que es generalmente incluida por la propia comunidad académica en sus prácticas docentes.

Los ODEs pueden ser usados por un estudiante que quiere aprender un determinado tema o por un profesor que quiere preparar algún material para su clase. Los usuarios pueden recuperar estos objetos por medio de búsquedas en repositorios. Estos objetos se almacenan utilizando metadatos descriptivos que proporcionan información adicional sobre los mismos. Existen distintos estándares de metadatos tales como Dublin Core (dublincore.org) y IEEE LOM (www.ieee.org), que utilizan categorías para describir no sólo el contenido del objeto (título, autor, palabras claves, idioma, etc.) sino también, como en el caso de LOM, permiten describir aspectos educacionales (nivel educativo, complejidad, etc.). La información almacenada en estos metadatos es fundamental para la mejor recuperación y recomendación de los ODEs. Sin embargo, en la mayoría de los casos, la información cargada en estos metadatos en los repositorios es de baja calidad o incompleta. Este grupo de I+D tiene trabajos previos respecto al desarrollo de sistemas que ayuden a la carga de ODEs mediante la extracción de metadatos [Pire et al., 2011; Casali et al., 2013].

La Universidad Nacional de Rosario ha creado en los últimos años un Repositorio Hipermedial institucional, RepHip (rephip.unr.edu.ar), para almacenar toda la producción académica, científica y de extensión de la misma. Este repositorio está implementado en la plataforma DSpace (dspace.org). Esta plataforma se ha adoptado para implementar el Sistema Nacional de Repositorios Digitales (SNRD) (repositorios.mincyt.gob.ar) de Argentina. Es necesario trabajar en distintas estrategias socio-culturales y políticas para la adopción por parte de la comunidad académica de estos repositorios, y trabajar también en el desarrollo de herramientas que ayuden en la usabilidad de estos repositorios [San Martín et al., 2013].

Uno de los objetivos de este grupo de investigación es desarrollar herramientas informáticas que soporten la gestión de estos repositorios en distintas facetas ayudando al administrador: en el proceso de recopilación de materiales plausibles de ser cargados en el repositorio; en el proceso de carga de ODEs mediante la extracción automática o semiautomática de metadatos de calidad, así como con mejoras propuestas en las interfaces de carga de los ODEs al repositorio. Otro de los objetivos es desarrollar herramientas informáticas que ayuden a los usuarios en la búsqueda mediante la recomendación de estos documentos.

Líneas de Investigación, Desarrollo e Innovación

En este proyecto se plantean líneas de investigación interrelacionadas.

La Red RIURE tiene por objetivo mejorar la usabilidad de los Repositorios Educativos, permitiendo búsquedas más intuitivas y sencillas. Para ello se

consideran distintos objetivos específicos:

- Introducir aspectos pedagógicos en los repositorios de material educativo para que sean tenidos en cuenta tanto en la catalogación como en la búsqueda, además de los metadatos técnicos.

- Introducir sistemas recomendadores basados en perfil de usuario, material y contexto, e integrable en los principales entornos de aprendizaje actuales.

- Utilizar estructuras semánticas como forma de describir los materiales, usuarios y contexto, de tal manera que se encuentren interrelacionados y permitan utilizar algoritmos más sofisticados de búsqueda y razonamiento. Esto también permitiría que el repositorio educacional se enlace con otros repositorios de información semánticamente anotada, ampliando la cantidad de recursos a los que tendrían acceso los usuarios.

En el proyecto [ING466] se propone contribuir al desarrollo de sistemas inteligentes basados en arquitecturas multiagentes, donde distintos tipos de agentes tienen que interactuar para asistir en la gestión de ODEs. Se plantea desarrollar sistemas de recolección de documentos en sitios restringidos para recopilar documentos plausibles de ser subidos a un repositorio, ayudando en sus tareas a los administradores de estas bibliotecas. Con sistemas de extracción automática de metadatos se puede asistir a los usuarios en el proceso de autoarchivo. Mediante sistemas recomendadores se puede asistir a los usuarios en el proceso de selección del material que mejor se adapte a sus preferencias, permitiendo búsquedas más intuitivas y eficientes, incorporando elementos de diálogo y negociación en este proceso.

En el proyecto [ING491] y en el proyecto UCA se trabaja en la expansión semántica de la búsqueda para producir

una estrategia de búsqueda temática, a partir de un conjunto de términos que ingresa el usuario que describen el tema de su interés. Además, se trabaja en cómo utilizar las características y preferencias del usuario para que un sistema recomendador no sólo recupere los recursos que respondan temáticamente a la consulta del usuario, sino que presente aquellos que sean los más adecuados de acuerdo a su perfil. Así, se ayuda al docente en la selección y preparación de material didáctico acorde a la temática a abordar, el tipo de material buscado, el nivel académico del curso y las características de cada grupo de alumnos [Casali et al., 2011].

Resultados y Objetivos

El objetivo de esta línea de investigación es el desarrollo de herramientas informáticas para ayudar a la gestión de repositorios de objetos de aprendizaje. Para esto se están desarrollando prototipos de sistemas para:

- dar soporte al administrador para las tareas de recopilación de documentos,
- facilitar el almacenamiento de documentos en repositorios con la extracción automática de metadatos,
- mejorar la búsqueda de información mediante expansión semántica de la consulta y sistemas recomendadores híbridos.

El desarrollo de herramientas de recopilación de información sirve para recopilar y clasificar información de sitios restringidos [Espinasse et al., 2009]. En esta dirección, se propuso en [Casali et al., 2014a] una arquitectura para automatizar la tarea de recopilación de documentos de texto dentro de un dominio web restringido con el objetivo de detectar objetos digitales educativos

plausibles de ser cargados en un repositorio institucional. Se ha implementado un prototipo que permitió evaluar la viabilidad de la arquitectura propuesta y experimentar cuál combinación de herramientas extractoras provee los mejores resultados en ciertos metadatos, tales como títulos, autores, filiaciones y emails de contacto. Así, para el prototipo se diseñó un extractor [Casali et al., 2014b] que combina herramientas de extracción, tales como ParsCit y Alchemy, y la utilización de expresiones regulares y bases de datos orientadas a grafos. Los resultados experimentales de este prototipo arrojaron buenos valores en la precisión y la cobertura de los metadatos mencionados. Las principales diferencias de esta propuesta con respecto a otros sistemas recopiladores son: la incorporación de la representación de la estructura de los sitios web correspondientes a las URLs semillas en una base de datos orientadas a grafos lo que permite recorrer y extraer información de nodos familiares enlazados, la diversidad de tipo de documentos a recopilar y la extracción de información tanto en idioma inglés como en español. Se espera que esta herramienta sea de utilidad a los administradores de repositorios institucionales ya que automatiza una parte importante de la tarea de recopilar documentos y poblar repositorios.

Por otro lado, en los últimos años el desarrollo de los sistemas recomendadores ha crecido notablemente siendo sistemas muy útiles en distintos campos y en particular en la educación. Estos sirven para recomendar actividades, estrategias, estudiantes con perfil similar a un usuario y en particular, objetos de aprendizaje. En esta rama de investigación también se aboca nuestra tarea, continuando con trabajos previos vinculados a la búsqueda

personalizada de objetos de aprendizaje [Casali et al., 2012] que explora y filtra las mejores opciones a partir de un perfil de usuario (preferencias y características) basándose en técnicas de filtrado basado en contenidos.

En [Casali et al., 2015] se han analizado sistemas recomendadores de objetos de aprendizaje desarrollados en la última década. De acuerdo a las distintas técnicas de recomendación utilizadas, se resalta que en su gran mayoría son sistemas híbridos donde se emplean dos o más tipos de filtrado para la recomendación. La mayoría de los sistemas analizados combinan un filtrado basado en contenidos con un filtrado colaborativo y algunos utilizan ontologías o reglas de producción para representar el conocimiento del dominio y mejorar la estrategia de recomendación. En experimentaciones realizadas, estos sistemas híbridos son los que demuestran mejores resultados respecto a los que usan una sola aproximación. En concordancia con estos resultados, se propone trabajar en Sistemas Recomendadores híbridos donde se puedan integrar distintas técnicas de recomendación como filtrado basado en contenidos, filtrado colaborativo y herramientas basadas en conocimiento. En este sentido, se trabajará para obtener un sistema recomendador de ODEs que personalice los resultados de la búsqueda utilizando el perfil del usuario, la valoración colaborativa de la comunidad y eventualmente, el modelo de diseño instruccional que se está siguiendo dentro de un ambiente de desarrollo de contenidos (LMS). También se explorará la recomendación basada en argumentación [Chesñevar et al., 2004], en el diálogo y en la crítica ([Ricci et al., 2007], [Chen et al., 2009]), para obtener una recomendación personalizada donde

las preferencias de los usuarios se puedan capturar en forma dinámica, en el proceso de interacción con los mismos.

Formación de Recursos Humanos

El equipo de trabajo está integrado por la Dra. Ana Casali, la Dra. Claudia Deco y la M. Sc. Cristina Bender. Dentro del marco de esta línea de investigación se desarrollan actualmente las siguientes tesinas de grado de la Licenciatura en Ciencias de la Computación de la UNR:

- Extracción automática de metadata de objetos de aprendizaje para su carga en repositorios. Santiago Fontanarrosa.
- Clasificación y búsqueda de documentos científicos, enriquecidas con una ontología. Fabricio Mahon.
- Agentes para la recolección de objetos digitales de aprendizaje en la Web. Santiago Beltramone.

Referencias

- [Casali et al., 2011] Casali A., Gerling V., Deco C. y Bender C. Sistema inteligente para la recomendación de objetos de aprendizaje. *Revista Generación Digital* Vol 9, No 1. pp. 88-95. Colombia. 2011.
- [Casali et al., 2012] Casali A., Deco C., Bender C. and Gerling V. Recommender System for Personalized Retrieval of Learning Objects. In *Educational Recommender Systems and Technologies: Practices and Challenges*, ERSAT (Eds. O. C. Santos and J. G. Boticario). aDeNu Research Group. UNED, Spain. , pp. 182-210; 2012.
- [Casali et al., 2013] Casali A., Deco C., Romano A., Tomé G.. An Assistant for Loading Learning Object Metadata: an Ontology Based Approach. *IJELLO* Volume 9, 2013.
- [Casali et al., 2014a] Casali, A. Deco C. Beltramone S. Automatic Gathering of Educational Digital Resources to Populate Repositories in *Proceedings of the XV International Conference on Human Computer*. INTERACCIÓN 2014. pp 495-501
- ACM 978-1-4503-2880-7/14/09. ACM Digital Library. Puerto de la Cruz, Tenerife, España. Septiembre 2014.
- [Casali et al., 2014b] Casali, A. Deco C., Bender C., Fontanarrosa S. Extracción Automática de Metadatos de Objetos Digitales Educativos. *Proceedings Novena Conferencia Latinoamericana de Objetos y Tecnologías de Aprendizaje LACLO 2014*, pp 23-29, Colombia. Octubre 2014.
- [Casali et al., 2015] Casali A., Deco C., Bender C., Cecchinell C., Ochoa X. Análisis de las distintas estrategias utilizadas en los sistemas recomendadores y su aplicación a recomendación de OAs *Reporte técnico de la Red 513RT0471* de Cytel. 2015.
- [Chen et al., 2009] Chen L., Pu P., Interaction design guidelines on critiquing-based recommender systems. *User Modeling and User-Adapted Interaction*. August 2009, Volume 19, Issue 3, pp 167-206.
- [Chesñevar et al., 2004] Chesñevar C. I., Maguitman A. G., Ricard G. A first approach to argument-based recommender systems based on defeasible logic programming. In *Proceedings 10th International Workshop on Non-Monotonic Reasoning (NMR 2004)*, Whistler, Canada, June. pp. 6-8. 2004.
- [Espinasse et al., 2009] B. Espinasse, S. Fournier, Fred Freitas. AGATHE: An Agent and OntologyBased System for Gathering Information about Restricted Web Domains, in: *IJEBR*, vol. 5, n° 3, pp. 14-34, 2009.
- [Pire et al., 2011] Pire T., Deco C., Casali A., Espinasse B. Extracción automática de metadatos de objetos de aprendizaje: Un estudio comparativo. En *Proc. VI Congreso de TE&ET*. Junio 2011. Salta, Argentina.
- [Ricci et al., 2007] Ricci F., Nhat Nguyen Q., Acquiring and Revising Preferences in a Critique-Based Mobile Recommender System, *IEEE Intelligent Systems*, vol.22, no. 3, pp. 22-29, May/June 2007, doi:10.1109/MIS.2007.43
- [San Martín et al., 2013] San Martín P., Bongiovanni P., Casali A., Deco C. Sociotechnological perspectives for Open Access Repositories development in the context of public universities in the central-eastern Argentina. *PKP Scholarly Publishing Conference*, DF Mexico. 2013.